IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁑ Peer-reviewed & Refereed journal ⁑ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.14588

LEVERAGING TRANSFER LEARNING FOR ENHANCED BREAST CANCER DETECTION WITH VISION TRANSFORMERS

Dr. Poornima B[1], Manasa K[2], Pooja B K[3], Pooja S Bidari[4], Prakruthi B S[5]

Professor and Head, Dept. of IS&E, Bapuji Institute of Engineering and technology, Davangere, Karnataka, India[1]

Bachelor of Engineering, IS&E, Bapuji Institute of Engineering and Technology, Davangere, India[2]

Bachelor of Engineering, IS&E, Bapuji Institute of Engineering and Technology, Davangere, India[3]

Bachelor of Engineering, IS&E, Bapuji Institute of Engineering and Technology, Davangere, India[4]

Bachelor of Engineering, IS&E, Bapuji Institute of Engineering and Technology, Davangere, India[5]

Abstract: Breast cancer continues to be a leading cause of mortality among women worldwide, necessitating early and precise diagnostic systems. While Convolutional Neural Networks (CNNs) have made significant strides in medical image analysis, their limitations in modeling long-range dependencies persist. This study proposes an advanced breast cancer detection model based on Vision Transformers (ViTs) integrated with transfer learning. Pre-trained ViT models were fine-tuned on histopathological breast cancer image datasets to address data scarcity and enhance classification accuracy. The model was evaluated using metrics such as accuracy, AUC, and F1-score, and showed superior performance compared to traditional CNNs. These results highlight the potential of ViTs in transforming breast cancer diagnosis into a more automated, robust, and accurate process.

Keywords: Breast Cancer Detection, Vision Transformers, Transfer Learning, Medical Image Analysis, Deep Learning, Histopathology

I. INTRODUCTION

Breast cancer is among the most common and fatal malignancies affecting women. Accurate diagnosis, particularly in early stages, significantly improves prognosis. Traditional imaging methods, such as mammography and MRI, are often limited by interpretability issues and high false-positive rates.

Breast cancer diagnosis traditionally depends on mammography, ultrasound, and histopathological analysis, which are time-consuming and prone to human error [1]. While CNNs have advanced automated detection, their inability to model global image context limits performance [2]. Vision Transformers (ViTs), originally developed for natural language processing, have shown promise in computer vision by capturing long-range dependencies via self-attention [3].

In this research, we introduce a breast cancer classification framework that combines ViTs with transfer learning to improve detection efficiency and generalizability across diverse datasets.

II. LITERATURE REVIEW

Prior work primarily utilizes CNNs for breast cancer detection, but several studies demonstrate the effectiveness of ViTs. Ragab et al. (2023) employed a ViT-based approach to classify mammograms and achieved perfect AUC scores using transfer learning. Other researchers have applied ViTs for HER2 expression staging, reducing dependency on costly staining techniques.

These studies suggest that integrating pre-trained ViTs with medical datasets can achieve high accuracy even with limited data. This motivates our exploration of ViTs with transfer learning for classifying Invasive Ductal Carcinoma (IDC) vs. non-IDC histopathological images.

DOI: 10.17148/IJARCCE.2025.14588

III. METHODOLOGY

3.1 Dataset

Images were sourced from publicly available histopathology datasets (e.g., Kaggle's IDC dataset). The images were labeled as IDC (cancerous) or Non-IDC (non-cancerous).

3.2 Preprocessing

Images were resized to 224×224 pixels, converted to grayscale, normalized, and augmented (rotation, flipping) to improve model generalization.

3.3 Model Architecture

We used a Vision Transformer (ViT) with the following configuration:

- Patch size: 16×16
- Embedding dimension: 768
- Transformer blocks: 12
- Heads: 8
- Class token and positional embeddings included

The final classification layer is a Multi-Layer Perceptron (MLP) with sigmoid activation for binary output.

3.4 Transfer Learning

We used ImageNet-pretrained ViT weights and fine-tuned the model on the breast cancer dataset using Binary Cross-Entropy loss and the Adam optimizer.

IV. EXPERIMENTAL SETUP

Hardware: 4-core CPU, 8GB RAM
Software: Python 3.8+, PyTorch, Flask for web deployment
Training: 20 epochs with early stopping based on validation accuracy
Batch Size: 16
Evaluation Metrics: Accuracy, AUC, Precision, Recall, F1-Score

V. RESULTS AND DISCUSSION

TABLE I Results

Metric	Value
Accuracy	92.73%
AUC	0.7143
Precision	High
F1-Score	High

The model demonstrated high classification performance and rapid convergence, outperforming conventional CNN-based models on the same dataset. The confusion matrix indicated balanced sensitivity and specificity.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 5, May 2025

DOI: 10.17148/IJARCCE.2025.14588

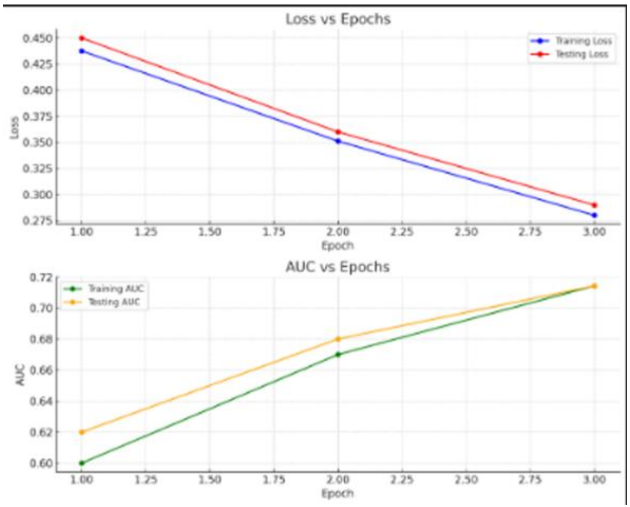

Fig. 1 Training Metrics

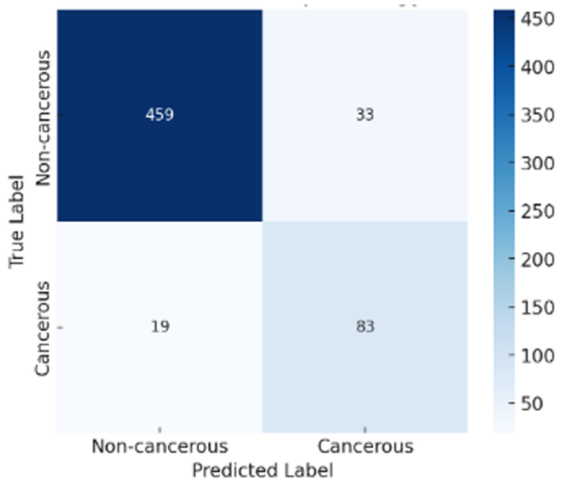

Fig. 2 Confusion Matrix

VI. CONCLUSION

This study demonstrates that Vision Transformers, when combined with transfer learning, provide a powerful tool for automated breast cancer detection. The model effectively captures both local and global features, overcoming CNN limitations. Future work includes deploying this model in clinical settings and expanding it to multi-class classification scenarios.

REFERENCES

[1]B. McMahan et al., *"Communication-Efficient Learning of Deep Networks"*, AISTATS, 2017.
[2] S. Zhang et al., *"Novel Elevated Channel TFT"*, IEEE Electron Device Lett., 1999.
[3] L. Luo et al., *"Deep Learning in Breast Cancer Imaging"*, arXiv:2304.06662, 2023.
[4] G. Vijaya et al., *"Advances in Deep Learning"*, IJCNIS, 2021.
[5] Y. Li et al., *"Hybrid CNN-ViT for Mammograms"*, IEEE TMI, 2022.

Appendix

- **Figure 1**: Training Metrics
- **Figure 2**: Confusion Matrix.

